GDPR and good data management

It’s easier than it sounds

Dr Lauren Cadwallader, Deputy Head of Scholarly Communication, Office of Scholarly Communication

Dr Rhys Morgan, Research Governance and Integrity Officer, Research Strategy Office
Good RDM is...

Research Data Management =

Organisation, storage, sharing and archiving of all data produced during research
And data is...?

Anything that supports the claims made in research

- Spreadsheets
- Code
- Images
- Videos
- Surveys
- Interviews
- Maps
- Physical samples
- Field notes...
And personal data is...

“Information relating to an identified or identifiable natural person” GDPR Article 4(1)

• I.e. any information about a living person that can be linked directly or indirectly to that person.

• Does not include data anonymised to the extent that a motivated third party could not reidentify the individual.

• Includes pseudonymised data (i.e. where identifiers are physically split from the data, but kept in the same organisation).
What does data protection legislation require of researchers?

The legislation imposes a range of requirements for use of personal data, but most research will be subject to an exemption:

• Research purposes – used where:
  • The standard provisions would seriously impair research
  • No damage or distress to data subjects
  • No individual decision-making about data subjects
  • Safeguards are in place

• Academic expression - used where:
  • Complying with the standard provisions would be incompatible with the academic purpose
  • There will be a publication in the public interest
What does data protection legislation require of researchers?

Good RDM will help you meet most of these requirements:

• Know and communicate your legal basis (research)
• Be transparent with data subjects (research)
• Process accurately and only what you need (research)
• Keep personal data secure (research)
• Process fairly, considering any ethical risks to the data subject (research)
• Comply with institutional accountability processes, e.g. ethical review (research and academic expression)
What is a DMP?

DMP = Data Management Plan

Plan outlining:
- Data to be collected during a project
- How it will be managed
- Covers the whole project ... and beyond

Usually 2-3 pages in length

Every project *should* have one
What is a DMP?

Contains information about:

- Sources of data
- Data input methods
- Back up strategies
- Plans for data sharing
- Ethical/legal restrictions
- Who is responsible for the data
What data are you going to collect?

Should be a simple question!

- Types of data
- Quantity of data
- File formats

Ensure that it is the right data!

- Only collect personal data if it is genuinely required and suitable for your research
- Unnecessary data cannot be held under either exemption

GDPR
Collecting personal data: Legal basis

Know your legal basis
• This will normally be that it is “necessary for the performance of a task carried out in the public interest”.
• **Not** normally consent.

Sensitive personal data
• E.g. race, political opinion, sex life, religious beliefs
• Must only be collected where proportionate to the risk (ethical approval usually required)
• Requires additional legal basis – “necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes”.

GDPR
Collecting personal data: being transparent

Need to inform data subjects about how their personal data will be used, must include:

• The legal basis you are using
• Who you *may* share the data with (with details of any sharing outside EEA)
• Any plans to combine the data with other datasets
• Retention period
• A link to the University’s privacy notice for research subjects
• Other information (e.g. risks of participation, withdrawal procedures, contact details of researchers should be provided for ethical and accountability reasons).
Storage & back-up

- Data should be stored in one location and backed up in 2 different locations
- Think about method, frequency, location and amount
- Consider your strategy for storage/back up in the field
- University provides storage services, which may be appropriate
Storage & back-up

- Personal data may only be used under the research exemption if appropriate security measures are taken.
- Ideally this will be prompt anonymization or pseudonymisation.
- Ensure that portable systems or devices are fully encrypted.
- Ensure physical security for hard copy data.
- See the [UIS website](https://www.uiswebsite.com) for appropriate University storage solutions.
Cloud storage

University provided clouds can store ‘Level 2’ data

<table>
<thead>
<tr>
<th></th>
<th>OneDrive for Business</th>
<th>Dropbox for Business</th>
<th>@cam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space included</td>
<td>1 TB</td>
<td>Unlimited</td>
<td>Unlimited</td>
</tr>
<tr>
<td>Price (per annum)</td>
<td>Free</td>
<td>£75 (+VAT)</td>
<td>Free</td>
</tr>
<tr>
<td>File history</td>
<td>90 days</td>
<td>30 days</td>
<td>Unlimited</td>
</tr>
<tr>
<td>Where are files stored?</td>
<td>Within UK</td>
<td>Within EU</td>
<td>Anywhere</td>
</tr>
</tbody>
</table>

GDPR
UIS Storage options

- Suitable for BIG data (>2Tb)
- Suitable for ‘Level 2’ data
- https://www.hpc.cam.ac.uk/research-data-storage-services

Research File Share
High-value data sets that require extra protection and are accessed frequently.

Research Data Store
Large active data sets, frequent access – for example, data processing on HPC.
Sharing with colleagues

- Think about what files you are sending your colleagues – do they really need the data?
- If sharing via email send an encrypted file
- Share anonymised data if possible
Sharing data publicly

- Required by most funders
- Might help with the Department’s REF return
- Data should be as open as possible and as closed as necessary
- Personal data should normally be anonymised unless consent for identifiable sharing has been given
- Remember to gain consent for open sharing when collecting your data
- UK Data Service runs a managed access repository if you can’t share data openly
Write a DMP. Plan properly. Help is available.
Research Data team

www.data.cam.ac.uk/DMPsupport

www.data.cam.ac.uk/funders
Research Data team

Andy Thwaites is your local Data Champion

First port of call for advice about data management

acgt2@cam.ac.uk
Research Ethics and Integrity

- Information on research ethics and integrity is available locally from ethics committees and departmental contacts.
- Central support is available from the Research Strategy Office:
  - researchethics@admin.cam.ac.uk
  - researchintegrity@admin.cam.ac.uk
  - www.research-integrity.admin.cam.ac.uk/
Thanks for listening!

Dr Lauren Cadwallader, Deputy Head of Scholarly Communication, Office of Scholarly Communication
info@data.cam.ac.uk

Dr Rhys Morgan, Research Governance and Integrity Officer, Research Strategy Office
rhys.morgan@admin.cam.ac.uk